Policy Research

WORKING PAPERS

Education and Employment

Population and Human Resources
Department
The World Bank
April 1992
WPS 889

On-the-Job Improvements in Teacher Competence

Policy Options and Their Effects on Teaching and Learning in Thailand

Stephen W. Raudenbush
Suwanna Eamsukkawat
Ikechuku Di-Ibor
Mohamed Kamali
and
Wimol Taoklam

Teacher supervision by effective principals is critical to improved teaching and learning in developing countries. Both teacher supervision and preservice training are far more important than inservice teacher training.

Policy Research

WORKING PAPERS

Education and Employment

WPS 889

This paper — a joint product of the World Bank's Education and Employment Division, Population and Human Resources Department, and Michigan State University - - is part of a larger effort in PHREE to understand differences in educational effectiveness. Support for this analysis was provided by the Bank's Research Support Budget for the research project On-the-Job Improvements in Teacher Competence: Policy Options and Their Effects on Teaching and Learning (RPO 676-36C). Support for data collection was provided by Project BRIDGES under a cooperative agreement between Harvard University and S&T ED, United States Agency for International Development (DPE 5824-A00-5076-00). Copies of this paper are available free from the World Bank, 1818 H Street NW, Washington, DC 20433. Please contact Cynthia Cristobal, room S6-033, extension 33640 (April 1992, 45 pages).

Teachers must hone their teaching skills on the job if the quality of primary education is to improve in developing countries. The five authors of this paper use a multi-level modeling procedure to examine two policy options for improving the competence of teachers already in the system: providing inservice training and encouraging regular classroom supervision.

After examining a nationwide sample of small rural primary schools in Thailand, they found that a teacher's experience in inservice training courses predicts neither instructional quality nor student achievement. In sharp contrast, intensity of supervision within a school significantly predicts both instructional quality and student achievement, after controlling for a variety of school, teacher, and classroom variables.

The effect of supervision is significant - - roughly the same as the effect of preservice education. Intensive field work in carefully selected rural schools suggests that supervision by effective principals is a critical component in a larger strategy to create and sustain an "ethos of improvement" in school teaching and learning.

On-the-Job Improvements in Teacher Competence: Policy Options and Their Effects on Teaching and Learning in Thailand

by
Stephen W. Raudenbush, Suwanna Eamsukkawat
Ikechuku Di-Ibor, Mohamed Kamali and
Wimol Taoklam

Background and Rationale

1. Improving teaching practice by improving teachers' skill and knowledge is essential if the quality of primary education in the developing world is to improve. This assertion is not intended to diminish the importance of investments aimed at improving facilities, developing coherent curricula, providing cost effective instructional materials, textbooks, or technology, or improving school management (Fuller, 1987; Lockheed, Vail & Fuller, 1987). But the impact of each of these investments on student learning depends upon the capacity of teachers to utilize resources effectively in classrooms (Lockheed and Komenan, 1989).

2. Policies aimed at boosting teacher competence are of central importance economically. In virtually all countries, teachers' salaries command the vast majority of educational expenditures. If policy options can be found that are relatively inexpensive, but that increase teacher competence, such options are likely to be highly cost effective.

3. Two historical developments help explain the emergence of teacher competence as a critical variable in the Third World educational policy-making. First, many countries have made substantial gains in providing schools and encouraging children to attend them. These increases in educational access have required the rapid expansion of school facilities, equipment, and materials, and the rapid credentialing and deployment of teachers. The focus in many regions of the developing world has begun to shift away from access and toward quality, and the recognition is widespread that many of the teachers now in place

lack the subject-matter knowledge and pedagogical skill required to stimulate an acceptable level of student learning (Fuller and Heyneman, 1989).

4. Second, comparatively little is known about cost effective approaches to increasing teacher competence. The contrast with what is known about educational provision is instructive. Associated with the growth in educational access is an associated growth in supporting technology. Sophisticated methods are available for forecasting growth in the demand for education, and these facilitate planning of school building and teacher hiring. Our state of knowledge regarding the selection, training, and supervision of those teachers, however, is comparatively primitive (Fuller & Snyder, 1991).

5. The focus of this paper is on policy options for improving teacher competence "on the job." Specifically, we compare effects of inservice training with the effects of supervision in small, rural Thai primary schools. By inservice training we mean courses of academic instruction that may or may not include components of classroom observation and feedback. Supervision refers to the activities of those invested with administrative authority over teachers to monitor, observe, evaluate, and provide feedback on classroom teaching. We note that inservice training and supervision may have elements in common. The primary difference is that inservice training courses are occasionally available and involve instruction outside of class. Supervision is a regular administrative responsibility. For example, in Thailand, principals are now expected to provide each teacher with regular classroom supervision either by visiting classrooms personally or by assigning expert teachers to do so.

We refer to supervision orchestrated by the school principal as "internal supervision." District supervisors are also charged with supervisory responsibilities that may involve classroom observation and feedback. We refer to this as "external supervision" because it is controlled at the district level.

6. Before describing our methods and results, we consider policy options for improving teacher competence on the job and describe the Thai context. We then review what is known about the effects of alternative strategies for increasing on-the-job learning in the developing world.

Importance of On-the-Job Learning

7. Both in developing and developed nations, researchers are intensively assessing alternative modes of preservice training. However, in many countries, it is the competence of teachers already on the job that must be enhanced if educational quality is to increase (Raudenbush, Bhumirat, & Kamali, in press). Many of the vast numbers of teachers hired during the period of rapid expansion of primary school access are still quite young, and where birth rates are declining, as in East Africa, teacher production is slowing. In many countries, available resources for substantial new hiring are now sharply curtailed. Under these circumstances, in which the teaching force is likely to change slowly during the next several years, policies that foster on-the-job learning are essential.

8. Discussions about policies to increase such on-the-job learning tend to focus exclusively on inservice training. And because the type, cost, and effects of alternative approaches to inservice training undoubtedly vary, this focus certainly has relevance to policy. However, alternative strategies for boosting learning on the job are available that may be more cost effective than inservice training. These include teacher supervision, teacher specialization, and the use of appropriate instructional materials.

9. Supervision. Observation followed by immediate and skilled feedback on classroom instruction provides a basis for teacher learning with direct relevance to teaching practice. In contrast, many inservice training schemes take teachers away from their classrooms and some divert thinking away from the problems that arise in their daily practice (Fullan, 1982; Wheeler, Raudenbush, & Pasigna, 1989).

10. Teacher specialization. In most countries, primary teachers are expected to be generalists, that is, to teach all areas of the curriculum to an intact class of children. However, in many schools, certain forms of specialization often arise informally in practice (Schmidt, 1990). For example, one teacher may teach math to children in several grades, while another teacher may concentrate on teaching reading to the same children. The arguments for specialization are that instruction will improve if teachers spend more time teaching the subject matter they know best; and that teachers will learn faster if they concentrate time and energy on those subjects in which they have a special interest. If these arguments are correct, policy makers may wish to encourage certain kinds of specialization

on a more widespread basis.

11. Use of instructional materials. The kinds of instructional materials available to teachers may influence the rate at which teachers learn subject matter and develop pedagogical skill. As Lockheed, Vail and Fuller (1986) have argued, texts and materials may provide teachers with a more structured and comprehensive representation of the subject matter than would otherwise be available to them. And involving teachers in the production, testing, and evaluation of new, locally-based materials may not only lead to use of appropriate materials; it may also have a significant educative effect on the teachers so involved (Wheeler et al., in press).

The Thai Context

12. Thai primary schools provide a useful context in which to study the effects of alternative policy options for improving teacher competence on the job. Having achieved near-universal primary access within the last 10-15 years, Thai policy makers have focused intense effort on the improvement of primary quality with special emphasis on improving the competence of practicing teachers.

13. From 1960 to 1980, the Thai government sought to universalize primary access, precisely during the period of most rapid population growth, requiring the very rapid certification and deployment of primary teachers. Then, in the late seventies, just as near-

universal access was being achieved, population growth slowed dramatically. The result has been that Thailand, like other middle-income countries, has had a relatively large reservoir of teachers, many lacking the needed subject matter knowledge and pedagogical skill. Because relatively few new teachers have been needed, reform of preservice education has provided limited opportunities to improve overall teacher competence. Thai policy makers have recognized this constraint, and therefore have focused considerable attention to improving the effectiveness of already-practicing teachers.

14. The reforms have taken remarkably diverse shapes (Wheeler et al., 1989). These include a wide variety of teacher inservice educational schemes; a national program of principal training with a strong emphasis on classroom supervision; and the granting of increased authority to school clusters, each having a resource center and a mandate to invoive teachers in developing and using innovative instructional materials. Teacher specialization has also arisen as a common practice, but not as a result of official policy (Schmidt, 1990). Rather, a number of teachers and principals have organized such specialization informally at the school level.

15. The circumstances that have impelled Thai policy makers to emphasize on-the-job learning are shared by other middle income societies but are not as clearly manifest in poorer developing nations, many of which continue to experience rapid population as they struggle to increase primary access. Nevertheless, each of these nations must be increasingly concerned with improving the competence of practicing teachers. The richness of Thailand's

experience in attempting to do so may be instructive to other nations contemplating policy options for fostering on-the-job teacher learning.

Prior Research on Learning on the Job

16. Research on teacher learning on the job has focused primarily on inservice training, and this research may be broadly classified into two categories: studies aiming to discover whether a program has positive effects, and studies aiming to identify which types of programs are most effective. Notably lacking are studies that compare inservice training to other policy options for improving teacher competence and student learning.

17. Does it work? The first type of study considers whether teacher participation in an inservice program has any demonstrable effect on teacher behavior or student learning. For example, Gage and Needels (1989) summarize results of 13 experiments in the US evaluating the effects of inservice teacher education programs. They found significantly positive effects on teaching practice in 12 cases and significant positive effects on student learning in 10 cases. Studies of this type typically have employed an experimental design contrasting two groups of teachers: a treatment group, which experienced the inservice training, and a control group, which did not. Nitsaisook and Anderson (1989) report an example of this type in Thailand, in which 100 primary teachers were assigned at random to an experimental or a control group. They found significant positive effects of training on student attitudes and achievement.

18. What works best? In the second type of research, the question is not whether training is better than no training, but rather, which types of training are most effective. Perhaps the dominant mode of research for this purpose is the literature review: The investigator collects a large number of studies of the first type (which compare particular training programs to control) and seeks to identify the characteristics of the programs producing the largest effects.

19. For example, Joyce and Showers (1980) reviewed 200 studies of a large variety of inservice training programs. They found that the most effective courses were characterized by five components: presentation of theory, "modeling" or demonstration of teaching techniques, practice teaching under simulated conditions, structured feedback after classroom observation, and coaching for application to new settings.

20. Ingvarson and Mackenzie (1988) reviewed the effects of week-long courses on computer education provided over a two-year period to 2000 teachers in Australia. They aimed to identify the factors that influenced teachers roles as "change agents" in disseminating knowledge in their schools after completion of the course. They found that teacher action as change agents varied from school to school and depended heavily on administrative support and follow-up assistance. They concluded that "Returns to investments in inservice education by school systems will be limited if planning for policy implementation goes no further than the provision of inservice courses alone, without the orchestration of follow-up support, both from within the school and from external sources."

(p 139).

21. In a similar vein, Wheeler et al.'s (1989) review of inservice education policy in Thailand criticizes "one-shot courses." They found that central government support, a focus on problems that arise in the practice of teaching, the active participation of teachers in the planning and implementation of the training, and sustained follow-up were required if inservice initiatives were to be utilized in teaching. Fullan (1982) summarized typical inadequacies of inservice courses in North America. These included short workshops based on topics not selected by the teachers involved with rare follow-up support or evaluation.

22. Not all attempts to identify "what works best" are based on reviews of past research. In some cases, investigators have designed studies in which more than two types of training are compared. Anderson and Djalil (1989) compared effects of "intensive training" and "minimal training" on Indonesian fifth grade teachers. They found that the intensively trained group changed their behavior more than did the minimally trained group and that their students also achieved significantly more in social studies.

23. The question of intensity is of obvious importance to policy because of the cost implications of intensive training. Gage and Grouws (1979) reported evidence indicating that minimally intensive and therefore inexpensive training can significantly improve teaching practice. Coladarci and Gage (1984) reported results casting doubt on this hypothesis. They speculated that training described as minimal may have produced positive results because the research designs involved frequent classroom observations that may have served the

function of monitoring and supervision.

24. Conclusions and implications for the present study. Studies we have identified, primarily though not exclusively conducted in developed nations, provide results that are broadly quite convergent. There is little reason to doubt that inservice teacher training programs can produce positive effects on teacher behavior and student achievement. The more important questions for policy involve the kinds of programs that work and the level of investment needed to produce results.

25. The literature seems to indicate that short-term courses without classroom follow-up are unpromising. Effective inservice instruction apparently requires classroom demonstrations, opportunities for teachers to practice and refine pedagogical techniques, and sustained follow-up, supported by classroom observation and feedback. Teacher involvement in the identification of course content and materials also may also be important. The effective models appear to be intensive and, therefore, expensive.

26. The strength of many of the studies we have cited is their experimental design, which facilitates causal inference. The weakness is that many of the programs investigated may be quite unlike the programs to which teachers are typically exposed. Indeed, the research provides little or no evidence on the summative effect of a nation's inservice education policy, and there is reason to suspect that inservice programs currently operating in the Third World rarely achieve the level of intensity required for success. Fuller (1987:281) has

commented on the lack of knowledge about effects of current inservice practice in the Third World:

> "Very little evidence exists on the effectiveness of in-service teacher training programs.... This scarcity of knowledge is in stark contrast to the increasing level of resources being invested in upgrading the skills of incumbent teachers. For instance, in the last decade, two thirds of the World Bank's education projects have included in-service training components. Only four studies have examined the influence of such efforts."

27. When researchers and policy makers seek to improve teachers' competence on the job, they apparently think primarily about inservice training. Notably lacking are studies that compare inservice training to other policy options for improving the competence of incumbent teachers. Yet the research on inservice training identifies the internal administrative support of the school as a critical ingredient to success. Sustained follow-up via classroom observation and feedback seem especially important in pedagogical reform. These are also critical ingredients in supervision.

28. The decision to emphasize inservice training rather than supervision has potentially important cost implications. Inservice training typically requires the temporary employment of instructors. Intensive inservice programs require extra spending to support classroom visitations. In contrast, internal supervision is conducted by those already permanently employed. The primary cost is the training of principals to conduct supervision, though some

approaches may impose the cost of periodically relieving expert teachers from teaching responsibilities.

29. The strength of the present study is that it provides summary evidence on the relative efficacy of both inservice education and supervision across a large population of small rural primary schools in an entire nation. The weaknesses are that the study is correlational, not experimental, so that causal inferences are made with caution; and that the evidence does not point to the types of programs that are particularly effective.

30. We shall consider as a key predictor variable the intensity of teachers' exposure to inservice training, as indicated by the duration of their training experiences and the content coverage of those courses. Similarly, we shall examine exposure to supervision, both internal and external. Outcome variables are the perceived quality of instruction and students' total achievement. We control a variety of covariates measured at the child, teacher and school levels. Effects of other policy options, including specialization and instructional materials, will be reported later.

Sample

31. Early in 1988, the Thai government, in collaboration of Project BRIDGES (Basic Research in Developing Educational Systems), conducted a representative national survey of approximately 400 primary schools, 4000 teachers, and 10,000 sixth-grade students. This

was a multi-purpose survey of conditions, practices, outcomes, and costs of primary schooling.

Design and Sample

32. The sampling design for the survey involved a multistage cluster sample with probability of selection proportional to size within the first two stages. There are 12 educational regions in Thailand, plus the Bangkok metropolitan area. Within each region, there are provinces (n=72); within provinces, there are districts and within districts there are schools.

33. The first stage of sampling involved a stratified random sample of 25% (n=18) of the provinces within strata constituted by educational regions. Within provinces, 20% of the districts were sampled at random and, within districts, 30% of the schools were sampled. Within each school every principal was administered the principal questionnaire. Moreover, within each school, every sixth grade teacher was selected; and one additional teacher from each other grade was sampled at random. Each teacher was administered the teacher questionnaire.

34. Finally, within schools, one sixth grade class was selected at random, although many schools contain only one sixth grade class. Within each class, all students were administered achievement tests in the five areas of the 1978 curriculum (Thai language, mathematics, life

experiences, work experience, and character development). Each student was also administered the student questionnaire. The resulting sample included 411 principals, 3808 teachers, and 9768 students.

35. A series of analyses of these data have already been reported (Raudenbush & Bhumirat, in press; Raudenbush, Bhumirat & Kamali, in press). However, a serious but unavoidable weakness in the survey design has prevented the integrated analysis pursued in the present paper. Specifically, it has been impossible to link each teachers' experience directly with the student ratings of instructional quality and with student measures of achievement. Our sub-sample was constructed to allow this linkage.

The Sub-sample

36. The survey procedures utilized by the Thai government allow the researcher to link school characteristics with individual teachers in each school and to link school characteristics with individual kids in that school. However, it has not been possible to link schools, teachers, and kids in a single analysis. Hence, the only teacher data available for predicting student outcomes have been aggregated levels of the teacher variables. Characteristic flaws with in this type of analysis are well known (c.f., Aitkin & Longford, 1986). The central flaw in the present case is that, to the extent teachers within a school vary on the independent variable, the use of the mean teacher value on that variable to predict student outcomes produces uninterpretable results.

37. Recently, however, the investigators have discovered that about 100 of the schools in the sample have only one sixth grade teacher. Thus, in these schools there is no ambiguity about the link between the teacher and the students. Moreover, these schools are of special interest to policy makers. These are small, rural schools, precisely those schools that provide the greatest challenge to policy aimed at improving teaching. Because of the availability of the nationally representative sample of schools, it is possible to locate these schools quite accurately in terms of the national distribution of a variety of important contextual variables.

38. Another advantage of using these schools is that, because our goal is to evaluate instruction during sixth grade rather than to estimate school effects over the entire course of primary schooling, it is now possible to utilize grade point average at fifth grade as a covariate for predicting student achievement. This is a powerful covariate with quite large differences between schools, probably because grades in Thailand are linked in part to regional and national examination scores.

39. After deletion of missing data, the obtained sample for the present analysis includes 103 sixth grade classrooms in 103 schools and a total of 2111 students. This compares to the national sample which has 411 schools, 3808 teachers, and 9768 students.

Instruments

40. Below we describe the outcome variables (student achievement and instructional quality), the key predictors (intensity of inservice training and supervision), and the key

covariates, that is student, teacher, and school characteristics that must be viewed as exogenous to the effects of inservice training and supervision.

41. The outcome variables and key predictors are described in detail below. The covariates are described briefly below, and in more detail in the appendix.

Outcome Variables

42. Student achievement. We restrict our attention to a measure of overall achievement derived from subtests developed and validated by the Office of the National Primary Education Commission. These include five subtests measuring Thai language achievement, six subtests measuring aspects of mathematics achievement, five subtests measuring achievement in life experiences (science and social studies), six subtests measuring achievement in character development and a subtest measuring work-oriented experience. The subtests were all significantly positively intercorrelated, each contributing to the overall internal consistency as indicated by Cronbach's alpha, which was .89.

43. Instructional quality. We asked students to estimate the frequency of occurrence of a variety of teaching behaviors or activities. To indicate the effectiveness of the teachers as leaders providing an organized academic environment, we asked the students how frequently their teachers explained the objectives of a new lesson, tested student knowledge after completion of a lesson, and provided feedback on the tests in order to clarify sources of

student errors. To indicate the success of the teacher in facilitating an active student role, we asked how frequently the student spoke in front of the class, how likely the student was to ask for explanations of unclear concepts, how frequently the teacher assigned students to prepare reports for presentation to the class, and how frequently the teacher allowed expression of student opinions on the conduct of a lesson. To assess the presence of fairness and personal concern, we asked students how likely their teachers were to inquire about reasons for student absence, whether the teachers were fair to every student, whether the students would approach their teachers privately about a problem, and whether the teachers provide special help for those who are behind. At the student level, this 12-item scale was found to have an internal consistency of .78. Despite the fact that the scale is made up of conceptually distinct sub-dimensions, each item in the scale contributed to the overall reliability of the scale, and a factor analysis suggested that a single global factor was tapped by these items. The reliability of the instructional quality variable aggregated to the school level is very high at .96[1].

Key Predictors

Inservice training. Teachers were asked first whether, during the past three years, they had received any kind of academic inservice training. Those responding affirmatively were then asked to state, for each instance of inservice training, the topic (curriculum and

[1]This high reliability of the school aggregate quality variable has two sources. First, there is substantial variation between classrooms on quality (58.3% of the variation is between classrooms; only 41.7% is between students within classrooms). This result suggests that there is substantial agreement among the students as "raters" of instructional quality within a classroom, a result that strengthens one's confidence in the instrument. Second, enough students per school/classroom were available (20 per classroom) to stabilize the aggregated classroom instructional quality score. If fewer students had been sampled, the reliability of the aggregate measure of quality would have been smaller.

content, teaching techniques, production and use of instructional materials, measurement and evaluation). In each case they were also asked to name the provider (staff within the school or cluster, district or regional offices, or teachers' college or university), and the duration (in days). Based on these responses, it was possible to construct, for each teacher, a measure of the total coverage (number of topics) and total duration of inservice training. For the present analysis, we restrict our attention to exposure to inservice training in terms of total duration.

44. Internal supervision. Teachers were asked first whether, during the past year, they had received any classroom supervision. Those responding affirmatively were then asked to state the central topic discussed during the feedback session (curriculum or content, teaching techniques, or measurement and evaluation). They were also asked the number of times they had received supervision. They were also asked about the provider of the supervision. We labelled as internal supervision that which is provided by the principal or by another teacher in the school. We labelled as external supervision that which was provided by a district or provincial supervisor.

45. In this analysis we consider effects of both internal and external supervision, and, for both types, we consider the effects of total duration (number of times) of supervision. Effects of topic will be reported later.

Covariates

46. Measures of student, community, and school background. Student level variables include SES, sex, linguistic background ("dialect"), nutrition, time needed to travel to school, pre-primary experience, prior repetition, and grade point average. School-level background variables included a scale measuring modernity of community infrastructure, a remoteness scale, mean student SES, and geographic region.

47. School resources. These included enrollment, pupil-teacher ratio, facilities and equipment, textbook availability, and availability of teaching materials. Enrollment is listed under school resources because prior research (c.f., Raudenbush & Bhumirat, in press) has indicated that larger schools tend to have a critical mass of resources not typically available in smaller schools. Other measures of resources, including scales measuring the physical condition of classrooms and school buildings, were not included because preliminary analysis showed no indication that these were related to the outcomes.

48. Teacher background. Teacher background variables exogenous to inservice training and supervision included preservice education (less than bachelors versus bachelors or more), sex, and principal age. Teacher age and experience were not included as predictors because preliminary analysis revealed that these were unrelated to the outcomes.

49. Transformations. Prior to their use in the analytic models, several variables were transformed into a logarithmic metric. These variables had positively skewed distributions, and scatterplots revealed that they tended to relate non-linearly to the outcomes. Variables transformed to a logarithmic metric included infrastructure, enrollment, pupil-teacher ratio, duration of inservice education, and duration of supervision[2].

Comparison of The National Sample and the Subsample

50. Descriptive statistics for selected variables are provided in Table 1. The table enables comparison between the obtained sample of small, rural primary schools with a single sixth grade classroom and a nationally representative sample of Thai primary schools. Table 1 shows that, in many ways, the schools in the sub-sample are disadvantaged relative to the schools in the population as a whole. They are substantially smaller and more remote, the infrastructure of the communities in which they are located is less modern, and fewer textbooks are available. They are likely to be located in the Northeast, the most impoverished section of Thailand. Their students are of lower SES, are less likely to speak the central Thai dialect, less likely to have experience pre-primary schooling, more likely to have repeated a grade, and they exhibit lower GPA's. Their principals are younger, which reflects the tendency of principals to move to larger schools as they are promoted. Their sixth-grade achievement levels are lower. Such disadvantaged schools are targets for improvement efforts according to Thailand's most recent national educational development plan.

---

[2]Linear relations between log-transformed variables and outcomes imply that relations between the original variable and the outcome are non-linear, suggesting a diminishing effect of the predictor as its value increases. For example, a positive effect of log-supervision implies that the benefit associated with an increment to supervision is greatest when little or no supervision is currently available and is smallest when high levels of supervision are already available (Raudenbush & Bhumirat, in press). Such interpretations are often substantively more reasonable than are interpretations implying that the value added to an outcome by investing in a resource is constant regardless of the current availability of that resource.

51. The teachers, however, are not disadvantaged. They have more education, on average, than do teachers in the nation as a whole, reflecting the policy of the Thai government to place new graduates in relatively remote, rural schools. They have, on average, experienced more inservice training, but are less likely to have received regular supervision.

Table 1: Descriptive Statistics for National Sample and Sub-Sample

Variable	National Sample		Sub-sample	
	m	sd	m	sd
(a) School Level				
Infrastructure	3.09	2.42	1.89	2.23
North	0.19	0.40	0.20	0.40
Central	0.18	0.39	0.16	0.37
South	0.13	0.12	0.07	0.25
Bangkok	0.15	0.13	0.08	0.27
Northeast	0.33	0.22	0.49	0.50
Pupil Teacher Ratio	21.32	6.78	20.49	7.93
Enrollment	425.23	515.91	164.77	85.15
Remoteness	7.77	2.70	8.49	2.23
Principal Age	44.78	8.75	41.13	8.50
Equpment and Facilities	0.00	0.19	0.00	0.42
Textbooks	26.26	8.41	20.49	7.93
Teaching materials	1.94	1.60	1.47	1.33
(b) Teacher Level				
Sex	0.37	0.48	0.40	0.49
Education	0.59	0.49	0.73	0.44
Inservice Training	1.61	0.89	1.84	0.73
Internal Supervision	9.16	11.19	7.57	7.51
(c) Student level				
Dialect	0.49	0.50	0.32	0.47
SES	0.00	0.69	-0.30	0.42
Breakfast	0.83	0.37	0.90	0.30
Time to School	18.77	15.08	16.48	12.60
Pre-primary education	0.49	0.50	0.38	0.48
GPA	4.81	1.39	4.49	1.41
Sex	0.50	0.50	0.49	0.50
Repetition	0.15	0.36	0.20	0.40
Instructional quality	45.00	7.24	44.51	9.33
Total Achievement	-0.12	11.83	-2.77	8.88

Analyses

52. The analysis proceeded through three stages for each of the two outcomes: a) specification of a model with no predictors, which yields baseline empirical evidence about the amount of variation at each level; b) specification of a model which identifies the relevant covariates; c) specification of models for the effects of each of the key predictors, controlling for the relevant covariates.

53. The analysis employed a hierarchical linear model as described by Raudenbush & Bryk, (1986) using the computer program of Bryk, Raudenbush, Congdon, and Seltzer (1988). Such models have also been labelled "multilevel models" (Aitkin & Longford, 1986; Goldstein, 1987). Lockheed and Longford (1991) were among the first to apply these techniques in studies of Third World countries, and their advantages in this context are also described in detail by Riddell (1989)[3]. The method enables separate specification of predictors measured at the student and the classroom level. Within-classroom variability is predicted by student-level variables; between-classroom variability is predicted by school-level characteristics[4].

---

[3]Other recent applications of these methods in research on developing nations are provided by Lockheed and Bruns (1990) Lockheed and Zhao (1991), Raudenbush, Kidchanapanish, and Kang (1991), Raudenbush and Bhumirat (in press), and Riddell and Nyagura (1991). The chief advantages of these approaches is that the analyst is able to take into account the effects of students' shared membership in classrooms and schools, increasing the validity of inferences about the effects of school and classroom policies, measured at the group level, on processes such as teaching and learning that occur within groups.

[4]The analysis at both levels is computed simultaneously by means of maximum likelihood. Estimation of residual variation within and between classrooms allows one to monitor the predictive power of the model at each level. The method is well-suited to studies of classroom effects because covariates measured at both levels can be controlled. This methodology overcomes the characteristic flaws of standard multiple regression methods in studies of school and classroom effects (see Raudenbush and Bryk's [1988] review). Such flaws include misestimated standard errors, aggregation bias, and erroneous assessments of variance explained.

54. In this analysis, students are viewed as nested within schools/classrooms. Because we shall be focusing on small rural primary schools with one classroom per sixth grade, school and classroom variation are indistinguishable.

55. Predictor variables at both levels were rescaled to have zero means so that intercepts would have meaning. However, indicator variables were not rescaled. Hence the intercept for each model may be interpreted as the expected score of an otherwise "typical" student having zero values on all indicators and attending a classroom having zero values on classroom-level indicators.

56. Missing data were treated with pairwise deletion at the student level and listwise deletion at the classroom level. The final sample therefore involved 103 classrooms and 103 schools. The distributions of the predictor variables in the complete sample (n = 125) and the reduced sample were compared and found to be highly similar.

Results

Predictors of Achievement

57. Base-only model. A remarkable proportion of the variation in total achievement was found to lie between classrooms. As Table 2 shows, the between -classroom variance with no predictors in the model was estimated to be 3785; the estimated within-classroom variance was 4100. Hence, 3785/(3785+4100) = .48 or 48% of the variation was between

classrooms, implying that classrooms vary substantially, either in their effectiveness or in their contextual and compositional characteristics.

Table 2: Predictors of Total Achievement

Predictors	Base-only Model			Final model		
	Coeff	SE	t	Coeff	SE	t
(a) School/Classroom Level						
Inercept	-24.45	6.25	-3.92	-50.29	11.55	-4.53
Infrastructure				37.25	15.80	2.36
Remoteness				6.53	2.76	2.37
Preservice Education				31.99	12.27	2.61
Internal Supervision				10.25	4.71	2.18
Inservice Training				(-5.23	6.68	-.78)[a]
(b) Student Level						
GPA				25.67	1.15	22.29
Repetition				-22.86	3.77	-6.06
SES				4.86	3.22	1.51
Time to School				-4.60	2.32	-1.98
Sex				-2.09	2.51	-0.83
Dialect				19.01	8.33	2.82
Breakfast				12.37	4.49	2.75
Variances						
Between class	3785			2693		
% Explained	0.0			28.8		
Within class	4099			3072		
% Explained	0.0			25.0		

[a]These are approximations based on regressing the residuals from the model on intensity of inservice training. Inservice training was also entered in the model and the results above confirmed.

58. Covariates and key predictors. Specification of covariates proceeded through four stages. First, the "best set" of student covariates were identified (SES and sex were maintained even if non-significant for purposes of comparison). Second, school composition variables were added to the model. Those significantly related to the outcome at a nominal ten percent level of significance were retained. Next, school resource indicators were included in the model, with significant predictors retained. Third, teacher background characteristics were included, with significant predictors retained. At each stage, residuals were regressed on predictors that had previously been dropped from the model to guarantee that the initial decision to drop a predictor remained correct in light of subsequent additions to the model.

59. Once a set of useful covariates had been identified, key predictors were added. Effects of intensity of inservice training and supervision were estimated both separately and jointly. Non-significant predictors were dropped in each case and the model re-estimated.

60. The rather painstaking procedure described above was necessary because the number of potential predictors was quite large relative to the number of classrooms. The goal was to achieve a theoretically plausible set of steps toward the identification of a parsimonious model. A large nominal significance level worked against Type II errors in this study having restricted degrees of freedom between classrooms.

61. Student-level predictors. Results for the final model appear in Table 2. Among the

student-level covariates, grade-five GPA was strongly and positively related to achievement, b = 25.67, t = 22.29; prior repetition was negatively related to achievement, b = -22.86, t = -6.06. Positive effects were found for speaking central Thai, b = 19.01, t = 2.82, and eating breakfast daily, b = 12.37, t = 2.75. Time needed to travel to school negatively predicted achievement, b = -4.60, t = -1.98. Once these variables were controlled, no significant effects of sex, SES, or pre-primary educational experience were present. Of course, each effect is adjusted for other effects in the model, and the importance of variables reflecting linguistic, social, nutritional, and pre-primary educational background are clearly underestimated, given that they are exogenous to GPA and repetition.

62. Standardized effect sizes for the dummy variables are easily computed. Given that the outcome has a standard deviation of 89, regression coefficients can be divided by 89 to obtain standardized mean differences between the group labelled one and the group labelled zero. The resulting standardized mean differences are -.26 for repetition, .21 for central Thai, and .14 for eating breakfast daily. Standardized regression coefficients for the continuous predictors may be computed by multiplying each raw regression coefficient by the ratio of the standard deviation of the predictor to the standard deviation of the outcome (Table 1). Results are .41 for GPA and -.04 for time to school. We note that the residual within-school variance is estimated to be 3072. Given the estimated within-school variance of 4100 based on the base-only model (i.e., the model having no predictors -- column 1 of table 2), we can infer that the student-level predictors account for (4100-3072)/4100 = .25 or about 25 percent of the within-school variation.

63. School-level predictors. Only four school-level predictors remained in the final model. It should be emphasized that other school-level variables have significant zero-order relations with the outcome, and some undoubtedly operate indirectly through those predictors in this model. For example, school-level background and resource variables may predict student GPA and internal supervision. Our present focus, however, is on the effects of policy-relevant predictors rather than identifying all pathways to achievement.

64. Results indicate that the community infrastructure (log metric) significantly and positively predicts achievement, b = 37.25, t = 2.36, as do preservice education, b = 31.99, t = 2.61 and remoteness, b = 6.53, t = 2.37. The positive effect of remoteness seems counter-intuitive, but is sensible when viewed as a selection artifact. Parents in less remote areas typically have choices among alternative schools, including private schools and larger public schools nearer town. Those remaining in a small public schools may be disadvantaged in other ways. In contrast, families in more remote schools likely have no options. This effect of remoteness is quite the opposite in the national sample which includes the much more advantaged large and urban schools.

65. Most notable is the significant positive effect of intensity of supervision (log metric), b = 10.25, t = 2.18. In contrast, nothing approaching a significant effect was found for any measure of inservice training, whether estimated with internal supervision included or excluded (see the "t-to-enter statistic" of -.78 associated with inservice training in Table 2).

66. The standardized effect size for pre-service education, the single dummy variable in the school-level model, may be computed by dividing its raw regression coefficient by the standard deviation of the outcome. However, the investigator has a choice of outcome standard deviation. One might choose the overall standard deviation of 89; or the school-level standard deviation of the outcome, 62 (square root of 3785 in column one of Table 2). This author prefers using the school-level standard deviation, because school-level variables can predict only school-level variation. The standardized effect in units of the school-level variation is .52, quite a substantial effect, implying that the expected achievement is .52 standard deviation units higher in classes taught by teachers with B.A.'s than in classes taught by teachers without B.A.'s, after adjusting for all other effects in the model. In units of the total standard deviation, this effect sizes is .36, still a substantial effect.

67. Standardized effect sizes for the other, continuously distributed predictors may be obtained by multiplying the raw regression coefficient by the ratio of the predictor to outcome standard deviations. Again, the investigator has a choice of standard deviation between the school-level standard deviation of 62 or the overall standard deviation of 89. Results for the three continuously measured predictors are given in Table 5. We add the standardized regression coefficient (as distinct from the standardized mean difference) for preservice education as well. Note that the standardized regression coefficients for preservice education (.23) and internal supervision (.20) are both non-trivial and of similar magnitude.

68. The residual variance between schools is 2693 (column two of table 2). This is a 29% reduction from the between school variance of 3785 asscciated with the base-only model (column one of Table 2).

69. Instructional quality as a predictor. To validate the study's measure of instructional quality, we specified a model in which the mean instructional quality of a classroom predicted total achievement, controlling for the relevant covariates. Results, presented in Table 3, column 1, indicate a positive effect of instructional quality, b = 1.72, t = 1.71. The effect is in the predicted direction, though the effect of instructional quality is actually smaller than the effect of internal supervision, possibly implying that the effect of internal supervision is mediated by teacher behaviors not captured in the instructional quality index. The standardized coefficient for instructional quality is .10 in units of the total standard deviation, and .15 in units of the between-classroom standard deviation. Note that in this case, availability of textbooks was added to the model because its "t-to-enter" indicated that it was needed. The implication is that, controlling for instructional quality, availability of textbooks becomes positively and significantly related to achievement. The standardized coefficient for textbook availability is .13 in units of the overall standard deviation and .18 in units of the between-classroom standard deviation.

70. To investigate whether the effect of internal supervision could plausibly be viewed as working through instructional quality, a model was specified in which both internal supervision and instructional quality were included as predictors, along with the necessary

covariates. Results (Table 3, column 2) show that the effect of internal supervision does reduce to insignificance when instructional quality is controlled, a result consistent with the hypothesis that the effect of internal supervision does indeed work through instructional quality to predict achievement.

Table 3: Instructional Quality as a Predictor of Total Achievement

Predictors	Without Int. Supervision			With Int. supervision		
	Coeff	SE	t	Coeff	SE	t
(a) School/Classroom Level						
Inercept	-40.59	11.54	-3.52	-44.27	11.84	-3.74
Infrastructure	25.15	15.64	1.61	30.32	16.06	1.89
Remoteness	6.65	2.74	2.43	6.62	2.73	2.43
Preservice Education	20.09	12.22	1.64	24.75	12.67	1.95
Internal Supervision				6.73	5.03	1.34
Textbooks	5.78	2.83	2.04	4.89	2.90	1.68
Instructional Quality	1.72	1.01	1.71	1.27	1.06	1.20
(b) Student Level						
GPA	25.55	1.15	22.17	25.59	1.15	22.21
Repetition	-22.82	3.77	-6.05	-22.77	3.77	-6.04
SES	4.83	3.22	1.50	4.82	3.22	1.50
Time to School	-4.56	2.32	-1.96	-4.54	2.32	-1.96
Sex	-2.09	2.51	-0.83	-2.10	2.51	-0.83
Dialect	15.28	8.31	1.84	16.69	8.37	1.99
Breakfast	12.34	4.49	2.75	12.27	4.49	2.73
Variances						
Between class	2644			2627		
% Explained	30.1			30.1		
Within class	3073			3073		
% Explained	25.0			25.0		

71. Effects of external supervision. The analyses described above were repeated for external supervision. In no case did the evidence suggest even a slight positive effect of external supervision.

Predictors of Instructional Quality

72. If either inservice training or supervision were to influence student achievement, they ought to do so through the mechanism of improving instructional quality. Our previous analysis had provided evidence that internal supervision does predict achievement, and it yielded some evidence that this effect worked at least in part through instructional quality as measured in this study. The implication is that internal supervision ought to predict instructional quality. We now turn to this question.

73. The analysis again proceeded through three stages. A "base-only" model enabled a partition of variation within and between classrooms. Next, we sought the "best" covariate set for predicting instructional quality. Finally, we estimated effects of internal supervision and inservice training as predictors of instructional quality.

74. Base-only model. The results indicated that 58.3% of the variation in this outcome was between classrooms; only 41.7% was between students within classrooms, implying that substantial agreement existed among students sharing a classroom regarding the quality of instruction in that classroom. Based on this level of agreement, and a sample size of about

20 students per classroom, the reliability of the classroom mean on instructional quality was .96 (see Rowan, Raudenbush, & Kang, 1991 for a discussion of reliability of aggregate indicators in multilevel analyses).

75. Because the classroom mean level of instructional quality was highly reliable and because the mean had been utilized as a predictor in the analysis of total achievement, we utilized mean instructional quality as the outcome in the subsequent analysis. Hence, the analysis was no longer multilevel, but rather was a single-level analysis at the classroom level.

76. Specification of the covariates. A series of regressions was estimated. First, school background indicators were used to predict instructional quality and the non-significant predictors (at a nominal 10 percent level) dropped from the model. Next, school resource variables were added, with non-significant predictors dropped and residuals regressed on excluded school background predictors to guarantee that no important covariate was ignored. Next, teacher background predictors were entered, with non-significant ones dropped and residuals again regressed on previously excluded predictors.

77. Specification of inservice training and internal supervision. Next, predictors indicating intensity of exposure to inservice training and internal supervision were added to the model, with non-significant predictors dropped and residuals again checked for possible associations with previously deleted variables. The results appear in Table 4. The model is remarkably

parsimonious, including just two predictors: preservice training, b = 2.36, t = 2.06, and internal supervision, b = 1.28, t = 2.99. These effects are equivalent to a classroom-level standardized regression coefficients of .20 and .29, respectively.

78. For comparative purposes, Table 4 also presents results when intensity of inservice training are added to the model. Results for preservice education and internal supervision are similar; note that the effect of inservice training fails to approach significance, b = 1.08, t = 1.04, though it is equivalent to a classroom-level standardized regression coefficient of .15 (the standard deviation of mean instructional quality was 5.26).

Table 4: Predictors of Instructional Quality

Predictors	Without Ins. Training			With Ins. Training		
	Coeff	SE	t	Coeff	SE	t
Intercept	-1.76	0.97	-1.81	-2.23	1.06	-2.10
Preservice Education	2.36	1.41	2.06	2.37	1.14	2.08
Internal Supervision	1.28	0.43	2.99	1.16	0.44	2.64
Inservice Training				1.12	1.03	1.09
Residual variance	25.46			25.42		
% Explained	9.7			10.8		

Table 5: Standardized Regression Coefficients[5]

Outcome	Predictor	Standardized Coefficient	
		In units of overall sd	In units of classroom sd
Total Achievement			
(From Table 2)			
	Infrastructure	.17	.25
	Remoteness	.16	.24
	Preservice Education	.16	.23
	Internal Supervision	.14	.20
(From Table 3)			
	Instruct. Quality	.10	.15
	Textbooks	.13	.18
Instructional Quality			
	Preservice Education	--	.20
	Internal Supervision	--	.29

[5]In each case, the same set of covariates is controlled and the predictor is considered by itself

Discussion

79. The need to boost the competence of the incumbent teacher is becoming increasingly apparent as Third World educational policy makers address the task of improving primary school quality. Especially in those countries where resources for hiring new teachers are scarce, or where the need for new teachers is diminishing, improving the effectiveness of teachers already on the job becomes paramount. When educators and researchers have considered this task, they have typically regarded inservice training schemes as the primary or even the sole policy option. Moreover, considerable research is now available indicating inservice training can significantly improve teacher performance and student outcomes, and the characteristics of successful and unsuccessful programs are gradually coming into focus. However, the present paper is based on two apparent weaknesses in research and thinking about improving the competence of teachers on the job.

80. First, few studies have sought to assess the summative effects of a society's investment in inservice training. Here the question is not whether inservice training can have an effect, though that question is important. Rather, the issue is whether empirical evidence can be found to indicate that a nation's investment in inservice training has produced a return.

81. If such evidence of the effect of inservice training is scanty, many would ask how inservice training programs can be improved. That is certainly an important question, and one that has absorbed the interest of numerous researchers, whose work we reviewed earlier. The results are quite coherent, and they indicate that such training programs must

be more intensive than they typically have been to produce demonstrable results. However, improving inservice training by increasing its intensity -- and cost -- is just one of several options. Hence, we have turned to a second question that has rarely been asked: What promise do policy options other than inservice training hold for improving teachers' on-the-job competence?

82. The findings of our study, which apply to small, primarily rural primary schools in Thailand mainly serving disadvantaged communities and students, are quite clear. There is little empirical evidence to indicate that a teacher's experience in inservice training courses predict improved instructional quality or student achievement. Nor is there evidence supporting the proposition that district supervisors' school visitations are helpful. In sharp contrast, there is clear evidence of a link between the intensity of internal supervision a teacher receives -- supervision provided by the principal or by designated teachers -- and the academic achievement of that teacher's students, after controlling for a variety of covariates measured at the school, teacher, and classroom level. There is equally strong evidence that their students view teachers receiving such supervision as providing higher quality instruction than teachers with less supervision. The evidence is consistent with a theory holding that the internal supervision improves student achievement by improving instructional quality. The magnitudes of effect of internal supervision on instructional quality and student achievement are roughly similar to the magnitudes of effect associated with preservice education (in terms of the contrast between a teacher's having a bachelor's degree versus not having one).

83. In interpreting the effect of internal supervision, the field research of Wheeler et al (in press), carried out in coordination with the survey reported here, is instructive. For purposes of that field work, four matched pairs of schools were identified from our survey data, with one matched pair in each of the four broad geographic regions of Thailand (the North, the South, the Northeast, and the Central areas). All were small, rural primary schools, and each pair was matched on community socioeconomic status, modernity of infrastructure, and linguistic background of the students. However, within each pair, schools were selected to be maximally different on total achievement. Teams of field workers assigned to each pair intensively interviewed staff and observed classes.

84. Their observations and interviews support our findings and elaborate their meaning. They found that in several of the most effective schools, regular internal supervision was a critically important component in the school principals' strategy to create and sustain a strong academic focus. Such schools were characterized by an "ethos of improvement" that encouraged teachers, for example, to come to school on time; to provide academic instruction in the afternoon as well as the morning; to use test results to evaluate instruction; and to discuss teaching in learning during lunch breaks. In some schools, principals who were effective at encouraging such an ethos were also active in mobilizing community resources to purchase instructional materials and in identifying resources at the district level that could support academic learning. In these settings it would be hard to imagine effective leadership without classroom supervision, but it is also clear that supervision is linked with a broader constellation of strategies for supporting the academic

mission of the school. Identifying appropriate inservice courses could certainly fit into such a constellation, but the field reports suggest that the locally-generated improvement efforts of which supervision is a crucial element stand in contrast to traditional approaches to inservice education, which too often emphasize a top-down orientation in which instructors "transmit" knowledge to the teachers. Though this sketch of the meaning of the "supervision effect" is somewhat speculative, it warns against a mechanical application of the practice of supervision divorced from the "ethos of improvement" that supervision can both reflect and reinforce.

85. The positive effect of internal supervision may also be viewed as consistent with past research, reviewed earlier, indicating that interventions into the lives of teachers may pay off when focused on problems of practice and viewed as useful in the eyes of the practitioner. In this crucial regard, good supervision may be similar to effective inservice education.

86. Regardless of which interpretation is most plausible, the results provide reason to doubt the proposition that many seem implicitly to embrace: that improving teacher competence on the job requires inservice training. At least one alternative strategy, support of regular internal supervision of instruction, should be considered as a potentially direct and cost efficient policy option for supporting effective instructional practice.

Appendix: Predictor Variables

Student, School and Community Background

1. Student socioeconomic status (SES) was derived from measures of father's education, mother's education, father's occupation, mother's occupation, and the natural logarithm of the amount of pocket money the child typically brings to school (alpha = .73), as reported by the students.

2. Students' linguistic background ("Dialect") was coded as "central Thai" vs "other," as reported by the students.

3. Student nutrition was measured by students' reports about how often they eat breakfast. An indicator variable was constructed: "Breakfast" (1 = daily, 0 = not daily).

3. Time needed to travel from home to school (in hours) was based on student reports.

4. Children's pre-primary experience was also coded dichotomously (1 = one or more years of pre-primary experience; 0 = no pre-primary experience) based on student reports.

5. GPA is student grade point average is a seven category variable (1 = 1.00 or lower; 2 = 1.01-1.50; 3 = 1.51-2.00, 4 = 2.01 - 2.50, 5 = 2.51-3.00, 6 = 3.01-3.50; 7 = 3.51-4.00).

6. Repetition is an indicator variable (0 = never repeated a grade; 1 = ever repeated a grade).

7. Sex is an indicator taking on a value of 0 if female, 1 if male. School level variables indicating community and school background were measured as follows:

1. The modernity of community infrastructure was indicated by a scale consisting of the sum of nine items including the presence in the local community of drinking water, paved roads, irrigation, telephone service, electricity, a hospital, a commercial bank, a factory, and a post office as reported by the principal. Each item was coded dichotomously; these dichotomies were then summed, and the resulting sum had an internal consistency of alpha = .81.

2. Student SES was aggregated to the school level to create the contextual variable "Mean SES".

3. Geographic regions were five: "North," "South," "Central," "Bangkok," and

"Northeast." Indicator variables were constructed for the first four of these categories so that the Northeast region constituted the reference group.

4. Remoteness is the mean of three items: distance to the district capital, distance to the market and distance to the highway (all in kilometers).

Availability of Resources

1. Enrollment is the count of students attending the school as reported by the principal.

2. Class size is the enrollment in the school's sole sixth grade classroom as reported by the principal.

3. Equipment was based on a 18 item scale including primarily equipment used for instruction ("hard technologies"), but also including some equipment that could be used for administration. These were quite highly intercorrelated and seemed best viewed as a single factor. Items included the presence or absence of a Thai typewriter, English typewriter, copying machine, slide projector, overhead projector, amplifier, radio cassette, radio, tape module, television, video cassette player, sewing machine, microscope, metronome, scale (for weighing), Thai musical instrument, international musical instrument, and water tank. The overall alpha was .77.

4. The availability of textbooks and workbooks, as indicated by the sum of the texts and workbooks available for student use across the five areas of the curriculum (alpha = .75: student report). Maximum total is 10.

5. Availability of instructional materials is based on 12 items indicating the teacher's assessment of each resources as sufficient or insufficient. Resources include a teaching manual, textbook, books to broaden the teacher's knowledge, paper, glue or staple, chalk, poster paper, charts, maps, science apparatus, radio, and bulletin board (alpha = .70).

References

Aitkin, M. & Longford, N. (1986). Statistical modeling issues in effectiveness studies. Journal of the Royal Statistical Society, B 149: 1-43.

Anderson, L. & Djalil, A. (1989). The impact of a research-based teacher training program on Indonesian teachers, classrooms, and students. Teaching and Teacher Education, 5,3, 165-178.

Bryk, A.S., Raudenbush, S.W. Seltzer, M., & Congdon, R. (1988). An Introduction to HLM: Computer Program and User's Guide. University of Chicago Department of Education.

Coladarci, T., & Gage, N.L. (1984). Effects of minimal intervention on teacher behavior and student achievement, American Educational Research Journal 21, 539-555.

Fullan, M. (1982), The Meaning of Educational Change. New York: Teachers College Press. Fuller, B. (1987). Raising school quality in developing countries: What investments improve school quality? Review of Educational Research, 57, 255-291.

Fuller, B. & Snyder, W., Jr. (1991). Vocal teachers, silent pupils? Life in Botswana classrooms. Comparative Education Review, 35,2.

Fuller, B. & Heyneman, S. (1989). Third world quality: current collapse, future potential. Educational Researcher, 18,2,12-19.

Gage, N.L. & Grouws, D.A. (1979). The Missouri mathematics effectiveness project: an experiment in fourth grade classrooms. Journal Educational Psychology, 71, 335-362.

Gage, N.L. & Needels, M.C. (1989). Process-product research on teaching: A review of criticisms. The Elementary School Journal, 89,3, 254-300.

Goldstein, H. (1987). Multilevel models in social and educational research. New York: Oxford University Press. Joyce, B. & Showers, B. (1980). Improving inservice training: Messages of the research. Educational Leadership, February, 379-385.

Ingvarson, L. & MacKenzie, D. (1988). Factors affecting the impact of inservice courses for teachers: Implications for policy. Teaching and Teacher Education, 2, 139-155.

Lockheed, M. & Bruns, B. (1990). School effects on achievement in secondary mathematics and Portuguese in Brazil. Policy, Research, and External Affairs Working Paper, Population and Human Resources Department. Washington, DC: The World Bank.

Lockheed, M. E., Vail, S., & Fuller, B. (1987). How textbooks affect achievement in

developing countries: Evidence from Thailand. Educational Evaluation and Policy Analysis, 8, 379-392.

Lockheed, M.E., & Komenan, A. (1989). Teaching quality and student achievement in Africa: The case of Nigeria and Swaziland. Teaching and Teacher Education, 5, 93-113.

Lockheed, M.E. & Longford, N. (1991). School effects in mathematics achievement gain. In S.W. Raudenbush and J.D. Willms (eds.) Schools, Pupils, and Classrooms: International Studies of Schooling from a Multilevel Perspective, pp 131-166. New York: Academic Press.

Lockheed, M. E. & Zhao, Q. (1991). The effects of local control on student achievement and attitudes in secondary schools in the Philippines. Paper presented at the annual meeting of the Comparative and International Education Society, Pittsburgh.

Nitsaisook, M. & Anderson, L.W. (1989). An experimental investigation of the effectiveness of inservice teacher education in Thailand. Teaching and Teacher Education, 5,4, 287-302.

Raudenbush, S.W., & Bhumirat, C. (in press). The distribution of resources for primary education and its consequences for educational achievement in Thailand. To appear in the International Journal of Educational Research.

Raudenbush, S.W., Bhumirat, C., & Kamali, M. (in press). Predictors and consequences of primary teachers' sense of efficacy and students' perceptions of the quality of instruction in Thailand. To appear in the International Journal of Educational Research.

Raudenbush, S. W. & Bryk (1986). A hierarchical model for studying school effects, Sociology of Education, 59,
1-17.

Raudenbush, S.W. & Bryk, A.S. (1988). Methodological advances in analyzing the effects of classrooms and schools on student learning. Review of Research in Education, 15, 423-474.

Raudenbush, S.W., Kidchanapanish, S., & Kang, S.J. (1991). The effect of pre-primary access and quality on educational achievement in Thailand. Comparative Education Review, 35,2,255-273.

Riddell, A. (1991). An alternative approach to the study of school effectiveness in third world countries. Comparative Education Review, 33, 481-497.

Riddell, A.R. & Nyagura, L.M. (1991). A multilevel analysis of school effectiveness in Zimbabwe. Paper presented at the annual meeting of the Comparative and International Education Society, Pittsburgh.

Rowan, B. Raudenbush, S.W., & Kang, S.J. (1991). School climate in secondary schools: A multilevel analysis. American Journal of Education, 99,2,238-266.

Schmidt, W.H. (1990). School-based, flexible staff specialization. Unpublished manuscript, College of Education, Michigan State University.

Wheeler, C.W., Chuaratanaphong, J., Bhumirat, C., Eamsukkawat, S., Pumsaard, S., Shinatrakool, R., Sirigirirakal, V., & Sookpokakit, B. (in press). School clusters in Thailand: A management strategy for improving primary school quality. International Journal of Educational Research.

Wheeler, C.W., Raudenbush, S.W., & Pasigna, A. (in 1989). Policy initiatives to improve primary school quality in Thailand: An essay on implementation, constraints, and opportunities for educational improvement. (BRIDGES Report Series No. 5). Cambridge, MA: Harvard Institute for International Development.

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS871	The Distribution of the Benefits from Social Services in Indonesia, 1978-87	Dominique van de Walle	March 1992	A. Bhalla 37699
WPS872	Romania's Evolving Legal Framework for Private Sector Development	Cheryl W. Gray Rebecca J. Hanson Peter G. Ianachkov	March 1992	CECSE 37188
WPS873	Measure and Interpretation of Effective Protection in the Presence of High Capital Costs. Evidence from India	Francois M. Ettori	March 1992	F. Ettori 32340
WPS874	The Trade Restrictiveness Index: An Application to Mexican Agriculture	James E. Anderson Geoffrey Bannister	March 1992	G. Ilogon 33732
WPS875	Rural Finance in Developing Countries	Jacob Yaron	March 1992	C. Spooner 30464
WPS876	Old Debts and New Beginnings: A Policy Choice in Transitional Socialist Economies	Ross Levine David Scott	March 1992	W Pitayatonakarn 37664
WPS877	Assessing Gains in Efficient Production Among China's Industrial Enterprises	Gary H. Jefferson Wenyi Xu	March 1992	CECSE 37188
WPS878	Adjustment and Private Investment in Kenya	Kazi M Matin Bernard Wasow	March 1992	D Ballantyne 37947
WPS879	Comprehensive Water Resources Management. A Concept Paper	Peter Rogers	March 1992	D. Ranger 31296
WPS880	Exchange Rate Policy, the Real Exchange Rate, and Inflation: Lessons from Latin America	Miguel A Kiguel	April 1992	R Luz 34303
WPS881	Dual and Multiple Exchange Rate Systems in Developing Countries. Some Empirical Evidence	Nita Ghei Miguel A. Kiguel	April 1992	R. Luz 34303
WPS882	Issues in Reforming Financial Systems in Eastern Europe The Case of Bulgaria	Alfredo Thorne	April 1992	L. Ly 37352

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS883	Malaria: The Impact of Treated Bed-Nets on Childhood Mortality in the Gambia	Pedro L. Alonso Allan G. Hill Patricia H. David Greg Fegan Joanna R M. Armstrong Andreas Francisco K. Cham Brian M Greenwood	April 1992	O. Nadora 31091
WPS884	Intercommodity Price Transmittal Analysis of Food Markets in Ghana	Harold Alderman	April 1992	C Spooner 30464
WPS885	The Impact of EC-92 on Developing Countries' Trade: A Dissenting View	A J Hughes Hallett	April 1992	A Kitson-Walters 33712
WPS886	Factors That Affect Short-Term Commercial Bank Lending to Developing Countries	Sudarshan Gooptu Maria Soledad Martinez Peria	April 1992	L Dionne 31426
WPS887	Power Sharing and Pollution Control Coordinating Policies Among Levels of Government	William Jack	April 1992	P Pender 37851
WPS888	Transformation of Agriculture in Central Eastern Europe and the Former USSR: Major Policy Issues and Perspectives	Csaba Csáki	April 1992	C Spooner 30464
WPS889	On-th-Job Improvements in Teacher Competence Policy Options and Their Effects on Teaching and Learning in Thailand	Stephen W. Raudenbush Suwanna Eamsukkawat Ikechuku Di-Ibor Monamed Kamali Wimol Tacklam	April 1992	C. Cristobal 33640